which is generated by said cluster analysis and defines a set of co-varying genes;

wherein the statistical significance of each of said sets of co-varying genes is determined by comparing the actual fractional improvement for the cluster defining said set to the distribution of fractional improvements for the cluster defining said set.

## REMARKS

Claims 1, 3-50, 58-64, 72-78, 89-100 and 105-124 are pending in the application. In the instant Amendment, claims 1, 18, 26, 29, 31, 38-39, 50, 64, 72, 96, 106 and 123 have been amended to clarify the present invention. Entry of the amendments presented herein is proper in that the amendments place the claims in condition for allowance or place the case in better condition for appeal. Upon entry of the above-made amendments, claims 1, 3-50, 58-64, 72-78, 89-100 and 105-124 will be pending. A marked version showing changes made to the amended claims is attached hereto as Exhibit A. A clean version of the pending claims, as amended, is attached hereto as Exhibit B.

Claim 1 has been amended to recite that the claimed method comprises determining, for each of a plurality of sets of cellular constituents in a plurality of response profiles, whether said set of cellular constituents is upregulated or downregulated by *each of* said first plurality of drug perturbations and that said consensus profile for said first plurality of drug perturbations *consists of* measurements of said set or sets of cellular constituents that are determined in said determining step to be upregulated or downregulated by *each of* said first plurality of drug perturbations (emphasis added). Claims 29, 38 and 39 have been amended similarly. Claim 31 has been amended accordingly. Support for the amendments is found in the specification at page 41, lines 21-24 and page 43, lines 4-5-31, lines 16-22, and lines 24-31. Claim 1 has also been amended to correct a typographical error.

Claim 18 has been amended to clarify that a fractional improvement is an improvement in total scatter with respect to *the center of said cluster* as compared *to total scatter with respect to the respective centers of the two clusters branching out of said cluster* (emphasis added). Claims 64, 96 and 123 have been amended similarly. Support for the amendments is found in the specification at page 29, lines 15-23. Claim 18 has also been amended to clarify that in the method fractional improvements are determined for each cluster

*which is generated by said cluster analysis and defines a set of co-varying cellular constituents* and that the statistical significance of each set is determined by comparing the actual fractional improvement for *the cluster defining the set* to the distribution of fractional improvements for *the cluster defining the set* (emphasis added). Claims 64, 96 and 123 have been amended similarly. Support for the amendments is found in the specification at page 27, lines 9-10 and lines 24-27; page 28, lines 1-9; and page 37, lines 17-19.

Claim 26 has been amended to clarify that a fractional improvement is an improvement in total scatter with respect to *the center of said cluster* as compared *to total scatter with respect to the respective centers of the two clusters branching out of said cluster* (emphasis added). Claims 50 and 106 have been amended similarly. Support for the amendments is found in the specification at page 37, lines 17-23; and at page 29, lines 15-23. Claim 26 has also been amended to clarify that in the method fractional improvements are determined for *each cluster which is generated by said cluster analysis and defines a set of response profiles* and that the statistical significance of each set is determined by comparing the actual fractional improvement for *the cluster defining the set* to the distribution of fractional improvements for *the cluster defining the set* (emphasis added). Claims 50 and 106 have been amended similarly. Support for the amendments is found in the specification at page 37, lines 17-23; page 27, lines 9-10 and lines 24-27; page 28, lines 1-9; and page 37, lines 17-19.

Claim 72 has been amended to make the language clearer.

No new matter has been added by the amendments. Entry of the foregoing amendments and consideration of the following remarks are respectfully requested.

## THE REJECTION UNDER 35 U.S.C. § 112, FIRST PARAGRAPH, SHOULD BE WITHDRAWN

Claims 18, 26, 50, 64, 96, 106, and 123 are rejected under 35 U.S.C. § 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the art that the inventor, at the time the application was filed, had possession of the claimed invention. Claims 14, 22, 47, 61, 92 and 119 are rejected under 35 U.S.C. § 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

With respect to claims 18, 26, 50, 64, 96, 106, and 123, the Examiner contends that the recitation that a "fractional improvement is an improvement in total scatter with respect to a cluster center in going from one cluster to two clusters" and the recitation of "the corresponding cluster" constitute new matter. Applicants have amended the claims to clarify that a fractional improvement is an improvement in total scatter with respect to *the center of said cluster* as compared *to total scatter with respect to the respective centers of the two clusters branching out of said cluster* (emphasis added). The rejection is therefore obviated. With respect to the rejection of the recitation of "the corresponding cluster," Applicants have amended the claims to recite that in the claimed methods fractional improvements are determined for each cluster which is generated by the cluster analysis and defines a set of, e.g., cellular constituents or response profiles, and that the statistical significance of each set is determined by comparing the actual fractional improvement for the cluster defining the set to the distribution of fractional improvements for the cluster defining the set. As such, the claims are not broader than the written support. The rejection is therefore obviated. Therefore, Applicants respectfully submit that the rejection of claims 18, 26, 50, 64, 96, 106, and 123 under 35 U.S.C. § 112, first paragraph, is obviated, and should be withdrawn.

With respect to claims 14, 22, 47, 61, 92 and 119, the Examiner contends that the algorithm of *hclust* is an essential subject matter for the practice of the above-listed claims and as such cannot be enabled by incorporation by reference to a printed publication. Applicants respectfully submit that the Examiner's contention is erroneous. "Essential material" is defined in the MPEP as that "which is necessary to (1) describe the claimed invention, (2) provide an enabling disclosure of the claimed invention, or describe the best mode" (see MPEP 608.01(p)). Thus the issue is whether the algorithm of *hclust* needs to be disclosed in order for the disclosure to enable the claimed invention, i.e., whether the algorithm of *hclust* is essential material under the definition in MPEP. Applicants reiterate that *hclust* is a well known algorithm for performing hierarchical cluster analysis and that software package S-Plus which includes *hclust* is a well known and widely used software package for performing statistical analysis. In the response to the previous Office Action dated October 9, 2001, Applicants provided evidence that S-plus and *hclust* algorithm are well known in the art by submitting as Exhibit C a printout of a review of S-plus entitled "A Flavour of S-Plus" by Bowman and as Exhibit D a printout of a review of S-plus entitled "S-

plus in Teaching" by Henery. Henery discloses that since 1989 S-plus was adopted as the official language for teaching all Statistics courses at University of Strathclyde, whereas Bowman discloses the use of S-plus as a teaching medium at University of Glasgow. Applicants also directed the Examiner's attention to Weinstein et al., 1997, Science 275:343-349, entitled "An information-intensive approach to the molecular pharmacology of cancer," already submitted as reference GO in the Information Disclosure Statement filed on October 5, 1999. As evidenced by note no. 21 on page 349, Weinstein uses S-plus in its cluster analysis calculations. Furthermore, Applicants submitted as Exhibit E printouts of web pages of Insightful Corp., a vendor of S-plus, demonstrating that S-plus is currently commercially available. As can be seen from these references, S-plus and *hclust* are indeed both well known and widely used in the art. Therefore, anyone of skill in the art would be readily capable of performing the claimed method of cluster analysis of response profile data using the S-plus software and the *hclust* algorithm. As such, Applicants respectfully submit that the disclosure of the algorithm of *hclust* is not necessary for a skilled artisan to make and use the claimed invention. The case law has consistantly held that what is well known is preferably omitted in the specification. A patent needs not teach, and preferably omits, what is well known in the art. *Spectra-Physics, Inc. v. Coherent, Inc.*, 827 F.2d 1524, 3 U.S.P.Q.2d 1737 (Fed. Cir. 1987). According to a more recent case, *Atmel Corp. v. Information Storage Devices, Inc.*, 198 F.3d 1374, 53 U.S.P.Q.2d 1225 (Fed. Cir. 1999):

> [35 U.S.C. §112] Paragraph 1 permits resort to material outside of the specification in order to satisfy the enablement portion of the statute because it makes no sense to encumber the specification of a patent with all the knowledge of the past concerning how to make and use the claimed invention. One skilled in the art knows how to make and use a bolt, a wheel, a gear, a transistor, or a known chemical starting material. The specification would be of enormous and unnecessary length if one had to literally reinvent and describe the wheel.

Therefore, Applicants respectfully submit that the rejection of claims 14, 22, 47, 61, 92 and 119 under 35 U.S.C. § 112, first paragraph, is in error, and should be withdrawn.

## THE REJECTION UNDER 35 U.S.C. § 112, SECOND PARAGRAPH, SHOULD BE WITHDRAWN

Claims 18, 26, 50, 64, 96, 106 and 123 are rejected under 35 U.S.C. § 112, second paragraph, as allegedly being indefinite. The Examiner contends that claims 18, 26, 50, 64, 96, 106 and 123 are vague and indefinite regarding step (a) because in the step a fractional improvement is determined but what is performed in order to obtain such an improvement is not indicated. In supporting his above-mentioned contention, the Examiner also contends that step (a) summarizes what is performed in steps (b)-(d).

Applicants respectfully point out that the rejection is based on an erroneous understanding of the term "fractional improvement." A fractional improvement is an "improvement" in the total scatter with respect to the center of a branch or cluster as compared to the total scatter with respect to the respective centers of the two lower branches or clusters which come out of the original brunch or cluster. For a definition of the term, Applicants respectfully direct the attention of the Examiner to the specification at page 29, lines 15-26. Thus, in one embodiment, the determination of a fractional improvement of unpermuted data simply involves carrying out a calculation according to Equation 13. Thus step (a) clearly recites what is performed in order to obtain a fractional improvement of the *unpermuted* data. Furthermore, Applicants also respectfully point out that step (a) does *not* summarize what is performed in steps (b)-(d). Indeed, step (b) performs permutation of the original data, step (c) performs cluster analysis on the permuted data, and step (d) performs determination of a fractional improvement of the *permuted* data. Therefore, Applicants respectfully submit that claims 18, 26, 50, 64, 96, 106 and 123 are not indefinite, and that the rejection under 35 U.S.C. § 112, second paragraph, should be withdrawn.

## THE REJECTIONS UNDER 35 U.S.C. § 103(a) SHOULD BE WITHDRAWN

Claims 1, 3-8, 10-13, 15-17, 19-21, 23-25, 27-46, 48, 49, 58-60, 62, 63, 72-78, 89-91, 93-95, 97-100, 105, 107-113, 115-118, 120-122 and 124 are rejected under 35 U.S.C. § 103(a) as being unpatentable over Eisen et al.,1998, *Proc. Natl. Acad. Sci. USA* 95:14863 ("Eisen"). Claims 1, 3-8, 10-13, 15-17, 19-21, 23-25, 27-46, 48, 49, 58-60, 62, 63, 72-78, 89-91, 93-95, 97-100, 105, 107-118, 120-122 and 124 are rejected under 35 U.S.C. § 103(a) as being unpatentable over Eisen in view of Welsh, U.S. Patent No. 5,686,114 ("Welsh"). The Examiner reiterates and maintains these rejections from the previous Office Action dated

October 9, 2001. Applicants respectfully disagree with the Examiner for reasons set forth below.

A finding of obviousness under 35 U.S.C. § 103(a) requires a determination that the differences between the claimed subject matter and the prior art are such that the subject matter as a whole would have been obvious to one of ordinary skill in the art at the time the invention was made. *Graham v. Deere*, 383, U.S. 1 (1956). The relevant inquiry is whether the prior art suggests the invention and whether the prior art provides one of ordinary skill in the art with a reasonable expectation of success. Both the suggestion and the reasonable expectation of success must be found in the prior art. *In re Vaeck*, 947 F.2d 488 (Fed. Cir. 1991).

Eisen teaches cluster analysis for analyzing the genome-wide expression data obtained from DNA microarray measurements. In Eisen, a microarray containing essentially every ORF from yeast is used to measure gene expression data of budding yeast during the diauxic shift, the mitotic cell division cycle, sporulation, and temperature and reducing shocks and a microarray with 9,800 cDNAs representing 8,600 distinct human transcripts is used to measure gene expression data of primary human fibroblasts stimulated with serum following serum starvation. The gene expression data obtained from the microarray measurement are then analyzed using cluster analysis to identify gene expression patterns. Eisen suggests that genes of similar function cluster together. However, although Eisen teaches that *genes* can co-vary and therefore can be clustered into co-varying sets, Eisen does not teach or suggest that some of such co-varying sets or clusters of genes, i.e., set or sets of genes, can be upregulated or downregulated by each of a particular collection of different drug perturbations, i.e., genes in each such set are either up-regulated or down-regulated by each of the collection of different drug perturbations. To support the rejection, the Examiner merely contends that Figures 1-3 "demonstrate data wherein different perturbations are analyzed wherein serum is added at different times" (see page 7 of the Office Action dated June 27, 2002). Firstly, Applicants respectfully point out that the perturbations in Eisen are not drug perturbations. Applicants further respectfully point out that merely displaying and analyzing data under different perturbations do not make the claimed invention obvious. In this regard, Applicants note that genes clustered in different sets are not in general co-varying under the different perturbations (that is why they are clustered in different clusters) and do not respond

to different perturbations similarly. For example, Eisen does not teach or suggest which set or sets of genes in Figures 1-3 are up-regulated or down-regulated by each of the perturbations. To make the claim language abundantly clear, Applicants have amended the claims to recite that the claimed method comprises determining, for each of a plurality of sets of cellular constituents in a plurality of response profiles, whether said set of cellular constituents is upregulated or downregulated by *each of* said first plurality of drug perturbations and that said consensus profile for said first plurality of drug perturbations *consists of* measurements of said set or sets of cellular constituents that are determined in said determining step to be upregulated or downregulated by *each of* said first plurality of drug perturbations (emphasis added). There is no teaching in Eisen with respect to such set or sets of genes or method of identifying such set or sets of genes. Nor does Eisen teach or suggest that the responses of such set or sets of genes that similarly respond to a particular collection of different drug perturbations can be used as the consensus profile for representing the response profiles of a cell in response to such a collection of drug perturbations. Therefore, Eisen's teaching that genes can be clustered into co-varying sets does not provide one of ordinary skill in the art with a suggestion and reasonable expectation of success that a group of the co-varying sets can have similar responses, i.e., upregulated or downregulated, to each of a collection of different drug perturbations and that such a group of the co-varying sets of cellular constituents can be identified and their responses used as the consensus profile of the cell in response to the collection of drug perturbations, i.e., to represent the response of a cell to the collection of drug perturbations. Thus, Eisen does not teach a method comprising determining, for each of a plurality of sets of cellular constituents in a plurality of response profiles, whether said set of cellular constituents is upregulated or downregulated by each of said first plurality of drug perturbations, each response profile in said plurality of response profiles (i) comprising measurements of a plurality of cellular constituents, and (ii) resulting from a different drug perturbation among said first plurality of drug perturbations to said type of cell or organism, wherein each set of cellular constituents in said plurality of sets of cellular constituents consists of cellular constituents that co-vary under a second plurality of perturbations or that are co-regulated, wherein said plurality of response profiles comprises at least five response profiles, and wherein said consensus profile for said first plurality of drug perturbations consists of measurements of said set or sets of cellular constituents that are

determined in said determining step to be upregulated or downregulated by each of said first plurality of drug perturbations.

The Examiner also contends that Applicants arguments presented in the response to the previous Office Action dated October 9, 2001 regarding projected response profiles is contrary to the factual basis. Applicants respectfully submit that the Examiner's contention is erroneous. In the response to the previous Office Action dated October 9, 2001, regarding the Examiner's contention that the usage of supervised clustering in Eisen is reasonably interpreted as the projected profiles of instant claims, e.g., claim 29, Applicants respectfully pointed out that the projection of response profiles according to a definition of cellular constituent sets is independent as to how the cellular constituent sets are defined and obtained and that the projection of response profiles is carried out after the cellular constituent sets have been obtained by, e.g., either supervised or unsupervised clustering. Therefore, since Eisen teaches only supervised clustering but not projection of response profiles, Applicants reiterate that Eisen teaches or suggests nothing about the projection of response profiles or projected response profiles.

The Examiner also contends that Eisen's teaching of correlation coefficient is a type of statistical significance. Applicants respectfully submit that the Examiner's contention is erroneous. To best illustrate the difference between correlation coefficient and statistical significance, Applicants respectfully direct the attention of the Examiner to page 25, lines 10-21, of the instant specification, which describes correlation coefficient and its use as a similarity metric; and to page 28, line 27 through page 33, line 16, which describes statistical significance and methods of its determination. From a comparison of the two sections, it is clear that correlation coefficient and statistical significance are two distinct concepts. Thus, by merely teaching correlation coefficient, Eisen teaches or suggests nothing about statistical significance. Therefore, Applicants reiterate that Eisen does not teach or suggest a method comprising a step of determining the statistical significance of obtained cellular constituent sets.

In addition, Applicants respectfully point out that Eisen does not teach or suggest a method for grouping measured response profiles comprising grouping response profiles among a plurality of response profiles into sets consisting of response profiles in which the responses of one or more sets of cellular constituents, e.g., genes, in each response profile are

similar among response profiles in the set. Eisen does not teach or suggest a method for determining the therapeutic efficacy of a drug comprising identifying one or more groups of sets of cellular constituents in one or more response profiles associated with exposure to the drug or drug candidate. Eisen does not teach or suggest a method of analyzing response data from a biological sample comprising grouping genes from the biological sample into sets of genes that co-vary in a plurality of at least five response profiles, and grouping the plurality of at least five response profiles into sets of response profiles that have similarly affected genes. Eisen does not teach or suggest a method of grouping sets of drug perturbations that similarly affect cellular constituents in a cell type or organism among a plurality of drug perturbations comprising grouping response profiles among a plurality of at least 5 response profiles in sets, each of said sets of response profiles consisting of response profiles in which the responses of one or more sets of cellular constituents are similar among the response profiles in the set.

With respect to the rejection based on Eisen in view of Welsh, Applicants reiterate that Welsh merely teaches pharmaceutical compositions comprising an inorganic pyrophosphate for use in the treatment of a disease associated with inappropriate or inadequate ATP-binding cassette (ABC) protein activity. Welsh does not teach or suggest what is missing in Eisen, i.e., that a group of the co-varying cellular constituent sets can have similar responses, i.e., upregulated or downregulated, to each of a collection of different drug perturbations and that such a group of the co-varying sets of cellular constituents can be identified and their responses used as the consensus profile of the cell in response to the collection of drug perturbations, i.e., to represent the response of a cell to the collection of drug perturbations. Welsh does not teach or suggest projection of response profiles onto basis cellular constituent sets or projected response profiles. Welsh does not teach or suggest a method comprising a step of determining the statistical significance of obtained cellular constituent sets. Nor does Welsh teach or suggest method for grouping measured response profiles comprising grouping response profiles among a plurality of response profiles into sets consisting of response profiles in which the responses of one or more sets of cellular constituents, e.g., genes, in each response profile are similar among response profiles in the set; a method for determining the therapeutic efficacy of a drug comprising identifying one or more groups of sets of cellular constituents in one or more response profiles associated with exposure to the drug or drug candidate; a method of analyzing response data from a biological

sample comprising grouping genes from the biological sample into sets of genes that co-vary in a plurality of at least five response profiles, and grouping the plurality of at least five response profiles into sets of response profiles that have similarly affected genes; or a method of grouping sets of drug perturbations that similarly affect cellular constituents in a cell type or organism among a plurality of drug perturbations comprising grouping response profiles among a plurality of at least 5 response profiles in sets, each of said sets of response profiles consisting of response profiles in which the responses of one or more sets of cellular constituents are similar among the response profiles in the set.

Therefore, Applicants respectfully submit that neither Eisen nor Eisen in view of Welsh renders the present invention obvious, and that the rejection under 35 U.S.C. § 103(a) over Eisen and the rejection under 35 U.S.C. § 103(a) over Eisen in view of Welsh should be withdrawn.

## CONCLUSION

Applicants respectfully request entry of the foregoing amendments and remarks into the file of the above-identified application. Applicants believe that each ground for rejection has been successfully overcome or obviated, and that all the pending claims are in condition for allowance. Withdrawal of the Examiner's rejections and allowance of the application are respectfully requested.

Respectfully submitted,

Date December 27, 2002

32,605

Adriane M. Antler (Reg. No.)

**PENNIE & EDMONDS** LLP
1155 Avenue of the Americas
New York, New York 10036-2711
(212) 790-9090

Enclosures

**EXHIBIT A: MARKED VERSION OF THE AMENDED CLAIMS**
U.S. APPLICATION SERIAL NO. 09/220,142
(ATTORNEY DOCKET NO. 9301-035-999)

(as amended December 27, 2002)

1. (Six Times Amended) A method of determining a consensus profile for a first plurality of drug perturbations to a cell type or organism, said method comprising determining, for each of a plurality of sets of cellular constituents in a plurality of response profiles, whether said set of cellular constituents is upregulated or downregulated by each of said first plurality of drug perturbations, each response profile in said plurality of response profiles (i) comprising measurements of a plurality of cellular constituents, and (ii) resulting from a different drug perturbation among said first plurality of drug perturbations to said type of cell or organism, wherein each set of cellular constituents in said plurality of sets of cellular constituents consists of cellular constituents that co-vary under a second plurality of perturbations or that are co-regulated, wherein said plurality of response profiles comprises at least five response profiles, and wherein said consensus profile for said first plurality of drug perturbations [comprises] consists of measurements of said set or sets of cellular constituents that are determined in said determining step to be upregulated or downregulated by each of said first plurality of drug perturbations.

18. (Three Times Amended) The method of claim 17, wherein said cluster analysis is carried out by a hierarchical clustering method, and wherein the objective statistical test comprises:

(a) determining for each cluster which is generated by said cluster analysis and defines a set of co-varying cellular constituents an actual fractional improvement in the cluster analysis of the cellular constituents based on the unpermuted responses of said cellular constituents, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster;

(b) generating permuted responses of cellular constituents by means of Monte Carlo randomization of perturbation index for the response of each cellular constituent across all perturbations;

(c) performing said cluster analysis on the permuted responses of cellular constituents;

(d) determining for each cluster which is generated in step (c) and defines a set of co-varying cellular constituents the fractional improvement in the cluster analysis of cellular constituents based on the permuted responses of cellular constituents, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster; and

(e) repeating steps (b) through (d) so that a distribution of fractional improvements in the cluster analysis of the cellular constituents is obtained for each said cluster which is generated by said cluster analysis and defines a set of co-varying cellular constituents;

wherein the statistical significance of each of said sets of co-varying cellular constituents is determined by comparing the actual fractional improvement for the [corresponding] cluster defining said set to the distribution of fractional improvements for the [corresponding] cluster defining said set.

26. (Three Times Amended) The method of claim 25, wherein said cluster analysis is carried out by a hierarchical clustering method, and wherein the objective statistical test comprises:

(a) determining for each cluster which is generated by said cluster analysis and defines a set of response profiles an actual fractional improvement in the cluster analysis of the unpermuted response profiles, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster;

(b) generating permuted response profiles by means of Monte Carlo randomization of cellular constituent index for each response profile across the measured cellular constituents;

(c) performing said cluster analysis on the permuted response profiles;

(d) determining for each cluster which is generated in step (c) and defines a set of response profiles the fractional improvement in the cluster analysis of the permuted response profiles, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster; and

(e) repeating steps (b) through (d) so that a distribution of fractional improvements in the cluster analysis of the response profiles is obtained for each said cluster which is generated by said cluster analysis and defines a set of response profiles;

wherein the statistical significance of each of said sets of response profiles is determined by comparing the actual fractional improvement for the [corresponding] cluster defining said set to the distribution of fractional improvements for the [corresponding] cluster defining said set.

29. (Four Times Amended) A method of determining a consensus profile for a first plurality of perturbations to a cell type or organism, said method comprising determining, for each of a plurality of sets of cellular constituents in a plurality of projected profiles, whether said set of cellular constituents is upregulated or downregulated by each of said first plurality of perturbations, each projected profile in said plurality of projected profiles

(i) resulting from a different perturbation among said first plurality of perturbations to said type of cell or organism, and

(ii) comprising measurements of a plurality of cellular constituents in said type of cell or organism that have been projected onto basis cellular constituent sets, said basis cellular constituent sets being defined by co-variation of measurements of cellular constituents under a second plurality of different perturbations, wherein said consensus profile for said first

plurality of perturbations [comprises] consists of projected measurements of said set or sets of cellular constituents that are determined in said determining step to be upregulated or downregulated by each of said first plurality of perturbations.

31. (Three Times Amended) The method of claim 29, wherein the consensus profile is the intersection of the sets of cellular constituents activated or de-activated by each of said first plurality of perturbations.

38. (Five Times Amended) A method of determining a consensus profile for a first plurality of perturbations to a cell type or organism, said method comprising determining, for each of a plurality of sets of genes in a plurality of response profiles, whether said set of genes is upregulated or downregulated by each of said first plurality of perturbations, each response profile in said plurality of response profiles (i) comprising measurements of transcript levels for a plurality of genes, and (ii) resulting from a different perturbation among said first plurality of perturbations to said type of cell or organism, wherein each set of genes in said plurality of sets of genes consists of genes having transcripts that co-vary under a second plurality of perturbations or that are co-regulated, and wherein said consensus profile for said first plurality of perturbations [comprises] consists of measurements of transcript levels for said set or sets of genes that are determined in said determining step to be upregulated or downregulated by each of said first plurality of perturbations.

39. (Three Times Amended) A method for comparing a biological response profile to a consensus profile, said consensus profile [comprising] consisting of projected measurements of one or more sets of cellular constituents, said one or more sets having been identified among a plurality of sets of cellular [consituents] constituents in a plurality of projected response profiles, each of said one or more sets of cellular constituents being upregulated or downregulated by [a] each of said first plurality of perturbations, each projected response profile in said plurality of projected response profiles

(i) resulting from a different perturbation to said type of cell or organism, and

(ii) comprising measurements of a plurality of cellular constituents in said type of cell or organism that have been projected onto basis cellular constituent sets, said basis cellular

constituent sets being defined by co-variation of measurements of cellular constituents under a second plurality of different perturbations, said method comprising:

(a) converting the biological response profile into a projected response profile by projecting measurements of cellular constituents in said biological response profile onto said basis cellular constituent sets; and

(b) determining the value of a similarity metric between the projected response profile and the consensus profile.

50. (Three Times Amended) The method of claim 49, wherein said cluster analysis is carried out by a hierarchical clustering method, and wherein the objective statistical test comprises:

(a) determining for each cluster which is generated by said cluster analysis and defines a set of response profiles an actual fractional improvement in the cluster analysis of the unpermuted response profiles, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster;

(b) generating permuted response profiles by means of Monte Carlo randomization of gene index for each response profile across the measured genes;

(c) performing said cluster analysis on the permuted response profiles;

(d) determining for each cluster which is generated in step (c) and defines a set of response profiles the fractional improvement in the cluster analysis of the permuted response profiles, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster; and

(e) repeating steps (b) through (d) so that a distribution of fractional improvements in the cluster analysis of the response profiles is obtained for

each cluster which is generated by said cluster analysis and defines a set of response profiles:

wherein the statistical significance of each of said sets of response profiles is determined by comparing the actual fractional improvement for the [corresponding] cluster defining said set to the distribution of fractional improvements for the [corresponding] cluster defining said set.

64. (Three Times Amended) The method of claim 63, wherein said cluster analysis is carried out by a hierarchical clustering method, and wherein the objective statistical test comprises:

(a) determining for each cluster which is generated by said cluster analysis and defines a set of cellular constituents an actual fractional improvement in the cluster analysis of cellular constituents based on the unpermuted responses of said cellular constituents, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster;

(b) generating permuted responses of cellular constituents by means of Monte Carlo randomization of the perturbation index for each cellular constituent across all perturbations;

(c) performing said cluster analysis on the permuted responses of cellular constituents;

(d) determining for each cluster which is generated in step (c) and defines a set of cellular constituents the fractional improvement in the cluster analysis of cellular constituents based on the permuted responses of cellular constituents, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster; and

(e) repeating steps (b) through (d) so that a distribution of fractional improvements in the cluster analysis of the cellular constituents is obtained for each cluster

which is generated by said cluster analysis and defines a set of cellular constituents;

wherein the statistical significance of each of said sets of cellular constituents is determined by comparing the actual fractional improvement for the [corresponding] cluster defining said set to the distribution of fractional improvements for the [corresponding] cluster defining said set.

72. (Four Times Amended) A method for analyzing response data from a biological sample comprising

(a) grouping [cellular constituents] genes from the biological sample into sets of genes that co-vary in a plurality of response profiles, each response profile in said plurality of response profiles (i) comprising measurements of transcript levels of a plurality of genes, and (ii) resulting from a different perturbation to said biological sample; and

(b) grouping the plurality of response profiles into sets of response profiles that have similarly [affect] affected genes,

wherein said plurality of response profiles comprises at least five response profiles.

96. (Twice Amended) The method of claim 95, wherein said cluster analysis is carried out by a hierarchical clustering method, and wherein the objective statistical test comprises:

(a) determining for each cluster which is generated by said cluster analysis and defines a set of co-varying cellular constituents an actual fractional improvement in the cluster analysis of the cellular constituents based on the unpermuted responses of said cellular constituents, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster;

(b) generating permuted responses of cellular constituents by means of Monte Carlo randomization of the perturbation index for response of each cellular constituent across the set of perturbations;

(c) performing said cluster analysis on the permuted responses of cellular constituents;

(d) determining for each cluster which is generated in step (c) and defines a set of co-varying cellular constituents the fractional improvement in the cluster analysis of cellular constituents based on the permuted response responses of cellular constituents, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster; and

(e) repeating steps (b) through (d) so that a distribution of fractional improvements in the cluster analysis of the cellular constituents is obtained for each cluster which is generated by said cluster analysis and defines a set of co-varying cellular constituents,

wherein the statistical significance of each of said sets of co-varying cellular constituents is determined by comparing the actual fractional improvement for the [corresponding] cluster defining said set to the distribution of fractional improvements for the [corresponding] cluster defining said set.

106. (Twice Amended) The method of claim 105, wherein said cluster analysis is carried out by a hierarchical clustering method, and wherein the objective statistical test comprises:

(a) determining for each cluster which is generated by said cluster analysis and defines a set of response profiles an actual fractional improvement in the cluster analysis of the unpermuted response profiles, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster;

(b) generating permuted response profiles by means of Monte Carlo randomization of cellular constituent index for each response profile across the measured cellular constituents;

(c) performing said cluster analysis on the permuted response profiles;

(d) determining for each cluster which is generated in step (c) and defines a set of response profiles the fractional improvement in the cluster analysis of the permuted response profiles, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster; and

(e) repeating steps (b) through (d) so that a distribution of fractional improvements in the cluster analysis of the response profiles is obtained for each cluster which is generated by said cluster analysis and defines a set of response profiles;

wherein the statistical significance of each of said sets of response profiles is determined by comparing the actual fractional improvement for the [corresponding] cluster defining said set to the distribution of fractional improvements for the [corresponding] cluster defining said set.

123. (Twice Amended) The method of claim 122, wherein said cluster analysis is carried out by a hierarchical clustering method, and wherein the objective statistical test comprises:

(a) determining for each cluster which is generated by said cluster analysis and defines a set of co-varying genes an actual fractional improvement in the cluster analysis of the genes based on the unpermuted responses of said genes, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster;

(b) generating permuted responses of genes by means of Monte Carlo randomization of perturbation index for the response of each gene across all perturbations;

(c) performing cluster analysis on the permuted responses of genes;

(d) determining for each cluster which is generated in step (c) and defines a set of co-varying genes the fractional improvement in the cluster analysis of genes based on the permuted responses of genes, wherein said fractional improvement is an improvement in total scatter with respect to [a] the center of said cluster [center in going from] as compared to total scatter with respect to the respective centers of the [one cluster to] two clusters branching out of said cluster; and

(e) repeating steps (b) through (d) so that a distribution of fractional improvements in the cluster analysis of the genes is obtained for each cluster which is generated by said cluster analysis and defines a set of co-varying genes;

wherein the statistical significance of each of said sets of co-varying genes is determined by comparing the actual fractional improvement for the [corresponding] cluster defining said set to the distribution of fractional improvements for the [corresponding] cluster defining said set.